Gestural Behavioral Implementation on a Humanoid Robotic Platform for Effective Social Interaction*

LaVonda Brown[1] and Ayanna M. Howard[2]

***Abstract*— The role of emotions in social scenarios is to provide an inherent mode of communication between two parties. When emotions are properly employed and understood, people are able to respond appropriately, which further enhances the social interaction. Ultimately, effective emotion execution in social settings has the capability to build rapport, improve engagement, optimize learning, provide comfort, and increase overall likability. In this paper, we discuss associating dominant emotions of effective social interaction to gestural behaviors on a humanoid robotic platform. Studies with 13 participants interacting with the robot show that by integrating key principles related to the characteristics of happy and sad emotions, the intended emotion is perceived across all participants with 95.19% and 94.23% sensitivity, respectively.**

I. INTRODUCTION

Emotions yield a natural form of communication, in that they can be shown visually through facial expressions, vocal expression, and actions/body movements. When certain emotions are integrated into social settings they have the capability to create a comfortable, welcoming environment for all parties. This, in effect, will increase a person's willingness to engage in the social interaction. In the realm of human-robot interaction (HRI), emotions have the capability to enhance scenarios involving education [1]–[4], physical therapy [5]–[7], play partners [8], [9], companions [9], [10], elderly care [11], and weight-loss [10].

One of the key uses of emotions in human-robot interaction scenarios is to build a bond between the two entities. Typically, this bonding mechanism can be enhanced by having the robot exhibit forms of empathy. Empathy is a key factor used to enforce socially supportive behaviors [12]. Smiling and showing sensitivity to the individual's emotions enhance the interpersonal relationship, which ultimately leads to increased enthusiasm and learning [12]. In [3], Saerbeck was able to implement empathy best by simply having a robotic agent smile (happy face) when a task was completed correctly and frown (sad face) when the task was completed incorrectly. This study proved that the appropriate expression of empathy in a social interaction scenario is best visualized through a happy-sad continuum as shown in the circumplex model of affect (Fig. 1) [13]. The vertical axis represents arousal, whereas the horizontal axis (happy-sad continuum) represents valence. The use of other emotions such as anger, surprise, and nervousness as feedback were shown not to be essential for active engagement.

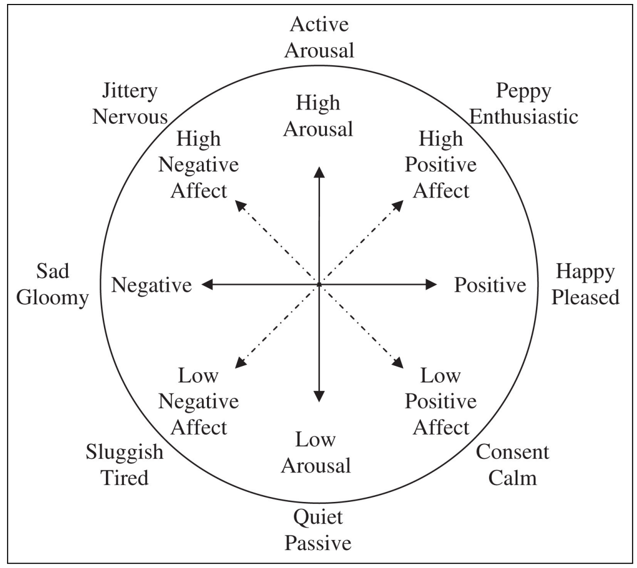

Fig. 1. The circumplex model of affect. [13]

Body movement can also be used to enable a robot to show forms of emotions in order to increase the quality of the interaction. During a case study involving a humanoid robot and children, [14] was able to analyze the effects of upward and downward head movement relative to positive and negative emotion. The humanoid robot was programmed to have six different base poses: anger, sadness, fear, pride, happiness, and excitement. Within each base pose, the head was positioned either up, down, or forward to make a total of 18 poses. The results showed that moving the head up improved the identification of pride, happiness, and excitement, while moving the head down improved the identification of anger and sadness. Fear was identified well regardless of the head's position. In general, moving the head up can enhance positive emotions, while moving the head down can enhance negative emotions.

In a similar study, Li and Chingnell analyzed how simple head and arm movements were able to communicate emotion in social robots [15]. Here, they used a teddy bear to implement various arm and/or head movements. They concluded that when head movements were compared to

*This research was supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1148903 and under Grant No. IIS-1208287. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

[1]LaVonda Brown is a Ph.D. student in the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA lavonda.brown@gatech.edu

[2]Ayanna M. Howard is a Motorola Foundation Professor in the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA ayanna.howard@ece.gatech.edu

arm movements, arm movements overall were perceived to be more lifelike. They also stated that these simple gestures alone do not provide a lot of information and recognition is low, which suggests that another medium to communicate emotion is needed.

Schegloff discussed the different effects achieved when changing the upper body parts versus the lower body parts [16]. He categorized the body into nine different stances: stance pose, hip pose, torso pose, shoulder pose, head pose, hip torque, torso torque, shoulder torque, and head torque. He concluded that lower body movement suggests “dominant involvement,” whereas upper body movement suggests “subordinate involvement.” This could possibly mean that lower body movements have more extreme effects on emotions, whereas upper body movements have less extreme effects. Although extreme emotions can be thought of as being less natural, in the realm of robotics, interaction is actually enhanced with the use of exaggerated motions [17]. In particular, Gielniak and Thomaz were able to present evidence that engagement is increased along with perceived entertainment value by over emphasizing movements during social interaction [17].

In an investigation involving the communication of musical expression through robotic gestures [18], robotic movements were derived from a perceptual test done by Dahl and Friberg [19]. Table I shows how the use of variables such as amount, speed, fluency, regularity, and direction on a mobile robotic platform are able to covey happiness, anger, and sadness. Burger also incorporated the work of [20] in this study, which stated that round shapes convey positive emotion and sharp/spiky objects oftentimes convey negative emotions. By using Table I, Burger had the robotic platform perform fluent, circular movements to convey happiness and jerky, sharp movements to convey anger [18].

In [4] we evaluated the positive effects of vocal expression during social interaction. Moving forward in this study, we examine implementation, perception, and potential impact of solely gestural behaviors in a similar setting. The primary limitation in the previous gestural studies is that they do not focus on the implementation of humanoid robotic gestures for effective social interaction – in [14], the only focus is the positioning of the head; in [15], [16], the gestures are implemented on a non-robotic platform; in [18]–[20], the gestures are developed to complement the various moods associated with music on a non-humanoid robotic platform. As such, in this paper, we focus on enhancing social interaction in HRI by ensuring that gestural behaviors are properly implemented on humanoid robotic platforms and understood by the individual. We detail a system that implements a range of gestures on a humanoid robotic agent for engagement and discuss the perception of the robot’s emotion from the human’s perspective. We want to test the hypothesis that by applying the key principles outlined in [14]–[16], [18]–[20] to a robotic social agent, individuals will be able to perceive the correct intended emotion. Section 2 discusses the robotic social agent and its implemented gestural behaviors. The experimental protocol used to evaluate the effectiveness of the robotic agent is presented in Section 3. The results and discussion points are made in Sections 4 and 5, and the conclusion is stated in Section 6.

TABLE I

IMPLEMENTATION OF THE ROBOT’S MOVEMENTS [18]

Movement Cue	**Happiness**	**Anger**	**Sadness**
Amount of Gesture	Large	Large	Small
Speed	Fast	Fast	Slow
Fluency	Fluent	Jerky	Fluent
Regularity	Regular, circular	Irregular	Regular
Direction of arm Movements	Upwards	Fast up and down	Slow up and down

II. GESTURAL BEHAVIORS

For this study, we derived a framework for implementing happy and sad emotions on a humanoid robotic platform. Anger can be detrimental to building rapport and establishing a level of comfort in social settings, so this emotion was not investigated. The major areas of interest when developing gestural behaviors are the head movements [14]–[16], arm movements [15], [16], [18], [19], and the overall size [16], [18], [19] and speed [18], [19] of the gesture. Based on prior studies, it is noted that moving the head in an upward position should convey happiness, while moving the head in a downward position should convey sadness [14]. Moving the arms in an upward position should convey happiness, while moving the arms slowly up and down should convey sadness [18]. The size (S) of the gesture is determined by the number of body parts in motion coupled with the range of motion of the movement as shown in (1) - (4).

$$S = f(A,B) \tag{1}$$

$$S_{large} = AB \tag{2}$$

$$S_{medium} = A\overline{B} + \overline{A}B = A \oplus B \tag{3}$$

$$S_{small} = \overline{A+B}, \tag{4}$$

where A is the number of active servos/joints and B is the range of motion. Based on this definition of “size,” a large gesture should convey happiness, while a small gesture should convey sadness [18]. The speed of the gesture is determined by the rate of change in the movement (not the total length of time). Based on this definition of “speed,” a fast gesture should convey happiness, while a slow gesture should convey sadness. The actual high/low thresholds associated with speed and size of the gestures were determined through empirical studies. Each of these key principles is outlined in Table II.

For the robotic social agent, we utilize the DARwIn-OP platform (Darwin) [21], a humanoid robot with 20 actuators, resulting in 6 DOF for each leg, 3 DOF for each arm, and 2 DOF for the neck (Fig. 2). To enable interaction with the

TABLE II

IMPLEMENTATION OF THE ROBOT'S MOVEMENTS

Key Principle	Parameter for Happy Emotion	Parameter for Sad Emotion
Head Direction	Upward	Downward
Arm Direction/Movement	Upward	Slow up & down
Gesture Size	Large	Small
Gesture Speed	Fast	Slow

human, Darwin was programmed with 15 gestural behaviors that were determined empirically. These behaviors were programmed using Darwin's default program ActionEditor in which we programmed a sequential set of actuator positions, with speed and timing constraints, to affect an appropriate gesture. A brief description of each gestural behavior is shown in Table III, and Fig. 2 displays an example of a happy, neutral, and sad gestural behavior.

Table IV outlines key attributes of how each gesture (G1 through G15) is associated with specific characteristics for depicting emotion. In particular we highlight the position of the head (upward or downward) [14], the movement of the neck, the direction/movement of the arms [18], the movement of the legs [16], and the overall size and speed of the gesture [18]. We also implement smooth, fluent, and regular movements for both the happy and sad gestures [18]. In addition, the purpose of highlighting the movements of each body part is to observe if dominant involvement (lower body movement) yield any significant differences when compared to subordinate involvement (upper body movement).

The "intended emotion" is defined as the emotion that should be perceived by the participant based on the key principles listed in Table II. When determining the intended emotion for each gesture, the sum of the happy principles was compared to the sum of the sad principles. For example, within Gesture 9, the attributes are head positioned downward, small gesture size, and a slow speed – because these are all are principles of the sad emotion, the resulting intended emotion is sad. In the cases where the gesture had principles of both happy and sad, the majority was chosen, and in the cases where the gesture had an equal amount of sad and happy principles, the intended gesture was classified as neutral. A more detailed analysis of each gesture and the number of happy/sad principles are shown in Table IV, along with its resulting intended emotion.

Our hypothesis states that by applying key principles outlined in Table II, individuals will be able to perceive the correct intended emotion (happy, neutral, or sad) implemented on a humanoid robotic platform. When the null hypothesis is accepted, the predicted intended emotion and the actual intended emotion will not be equivalent and/or the sensitivity of the intended emotion will be less than 75%. When the null hypothesis is rejected, the predicted intended emotion and the actual intended emotion will be equivalent and the sensitivity of the intended emotion will be greater than 75%. Sensitivity is the true positive rate (*TPR*) and (5) will be used to test the hypothesis.

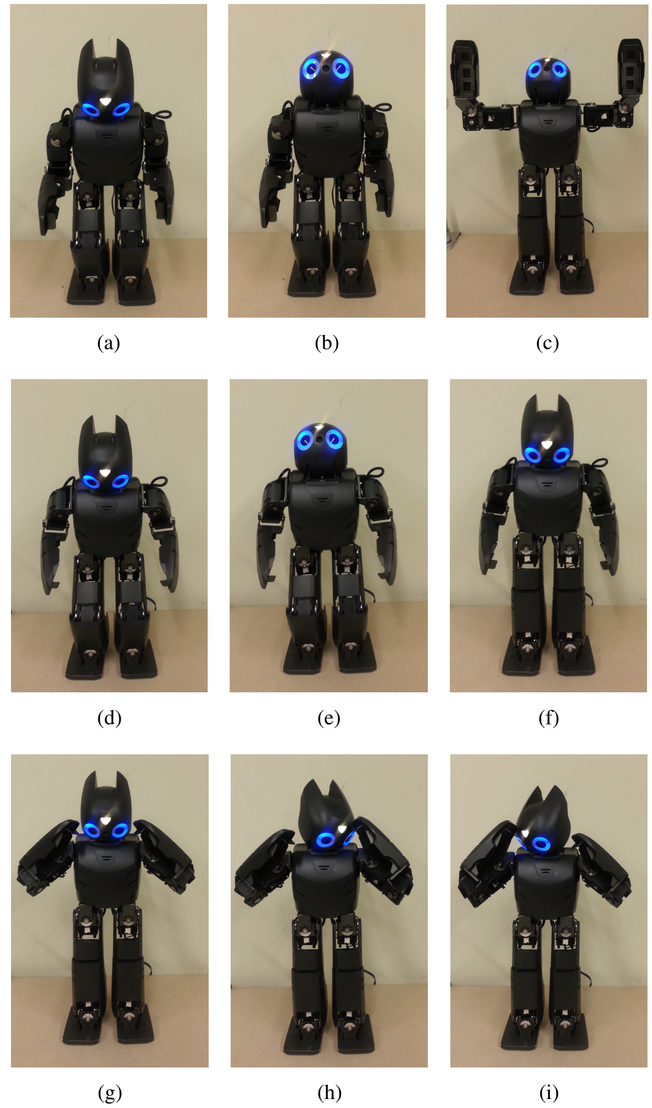

Fig. 2. G6 (happy) is broken down into three parts - (a) (b) (c). G5 (neutral) is broken down into three parts - (d) (e) (f). G12 (sad) is broken down into three parts - (g) (h) (i). Refer to Table III for each gesture's description.

$$TPR = TP/P = TP/(TP+FN), \tag{5}$$

where P is the number of positive instances, TP is the number of true positives, and FN is the number of false negatives.

III. EXPERIMENTAL DESIGN

To evaluate the perception of gestural behaviors implemented on the robotic social agent, we employed a single group design for this study. A total of 13 human participants took part in this experiment and all were recruited students from undergraduate and graduate studies at Georgia Institute of Technology in Atlanta, GA. The population consisted of both females and males in the age range of 18-34 years old (mean = 25.8, standard deviation = 3.9; Male: 8, Female: 5; undergraduate: 1, graduate: 12). During the study, the participant sat at a desk where Darwin stood approximately 2 feet in front of him or her. Once the participant was

TABLE III

DESCRIPTION OF GESTURAL BEHAVIORS FROM THE ROBOTIC EDUCATIONAL AGENT

Gesture	Description
G1	This gesture imitates the "field goal" symbol that society uses to communicate the idea of a job well done (in sports, this means that the team has scored). Darwin raises both of his arms simultaneously and forms a 90-degree angle with respect to the ground while his head is upward.
G2	Darwin bends his knees first, then straightens his legs while raising both arms to form a 90-degree angle with the ground and looking upward.
G3	Darwin simply moves his head only in an up and down motion.
G4	Darwin moves his head only in an up and down motion while raising both arms to form a 90-degree angle with the ground.
G5	Darwin bends his knees then straightens his legs repeatedly while moving his head up and down simultaneously.
G6	Darwin bends his knees first, then straightens his legs while raising both arms to form a 90-degree angle with the ground. Throughout this entire gesture, Darwin is moving his head up and down.
G7	Darwin bends his knees first, then straightens his legs while raising both arms to form a 90-degree angle with the ground. Throughout this entire gesture, Darwin is moving his head up and down (very subtle differences from G6).
G8	Darwin raises his left arm in a 90-degree agree towards his head. He then moves his arm up and down next to his face, as if he were scratching his head.
G9	Darwin simply lowers his head to the ground.
G10	Darwin lowers his head to the ground and then raises his hands to his head, as if they were holding his head
G11	Darwin lowers his head to the ground and then slowly shakes his head from side to side.
G12	Darwin lowers his head to the ground and raises his hands to his head, as if they were holding his head. He then slowly shakes his head from side to side.
G13	Darwin looks upward while raising his arms in the air and bringing them together, as if he were clapping.
G14	Darwin raises his left arm in a 90-degree agree, then pulls it down at a rapid speed.
G15	Darwin moves nods his head up and down while simultaneously moving his arms back in forth, as if he were engaging someone in conversation.

TABLE IV

KEY ATTRIBUTES OF EACH GESTURAL BEHAVIOR

Gesture	Head	Neck	Arms	Legs	Size	Speed	Happy	Sad	Intended Emotion
G1	Upward	–	Upward	–	Medium	Fast	3	0	Happy
G2	Upward	–	Upward	Bend	Large	Fast	4	0	Happy
G3	–	Up & Down	–	–	Small	Moderate	1	0	Happy
G4	–	Up & Down	Upward	–	Medium	Fast	2	0	Happy
G5	–	Up &Down	–	Bend	Medium	Moderate	0	0	Neutral
G6	–	Up & Down	Upward	Bend	Large	Fast	3	0	Happy
G7	–	Up & Down	Upward	Bend	Large	Fast	3	0	Happy
G8	–	–	Up & Down	–	Medium	Moderate	0	0	Neutral
G9	Downward	–	–	–	Small	Slow	0	3	Sad
G10	Downward	–	Midway	–	Medium	Slow	0	2	Sad
G11	Downward	Side to Side	–	–	Medium	Slow	0	2	Sad
G12	Downward	Side to Side	Midway	–	Large	Slow	1	2	Sad
G13	Upward	–	Upward	–	Medium	Fast	3	0	Happy
G14	–	–	Up & Down	–	Medium	Fast	1	0	Happy
G15	–	Up, Down & Side to Side	Midway	–	Medium	Moderate	0	0	Neutral

positioned, Darwin performed a gesture at random (Table III), and then returned to a standing rest position. The gestures were selected at random to ensure that order of the gestures presented did not have an effect on perception. If the participant did not see a gesture fully, or asked to view it again, Darwin would perform it again until the participant was ready to move forward to the next gesture. At the completion of each gesture, the participant recorded their perception of Darwin's behavior on a 5-point Likert scale (very happy-very sad) regarding their interaction with the robotic agent. This is repeated until all 15 gestures had been performed by Darwin and evaluated by the participant. The study was completed in approximately 10 minutes.

IV. RESULTS

To prove or disprove the hypothesis that the perception of emotion implemented on a robotic social agent can be determined by the key principles outlined in Table II, we analyze the results of the Likert scale and confusion matrices. First we look at the results of a 5-point Likert scale, where 1 is "very happy," 2 is "happy," 3 is "neutral," 4 is "sad," and 5 is "very sad." These results are shown in Fig. 3.

Table V depicts the confusion matrix associated with the results; the gestures are combined together based on the intended emotion happy, neutral, and sad. There were a total 104 instances of happy gestures, 39 instances of neutral gestures, and 52 instances of sad gestures. Because the data set is unbalanced (the number of instances vary for each emotion), we evaluate the sensitivity and the specificity of each emotion in Table VI. Sensitivity is the true positive rate (TPR), whereas specificity is the true negative rate (TNR).

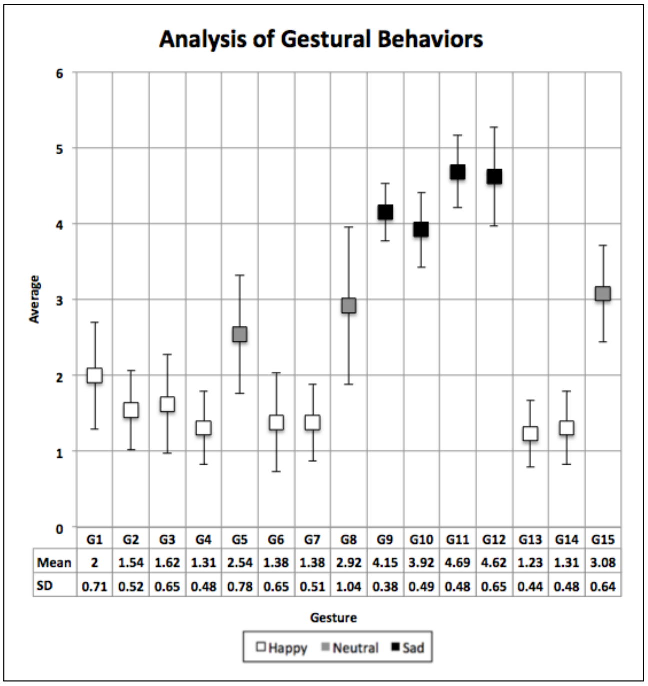

Fig. 3. Each gesture's average intended emotion is shown along with its associated mean from the 5-level Likert Scale. The upper and lower error bars are equivalent to one standard deviation.

V. DISCUSSION

At first glance, the standard deviations in Fig. 3 were all less than 0.8 except for one, and it was a common trend that the "neutral" emotions had higher standard deviations than the "happy" and "sad" emotions. An explanation for this is that the lack of dominant characteristics in the gesture caused confusion for the participants. Table V shows that G5, G8, and G15 all have 0 sad principles and 0 happy principles, so it is not a surprise that participants were confused with these gestures. Even during the actual testing, these gestures where oftentimes asked to be repeated for clarification. Therefore, this suggests that there must be a distinguishable amount of happy/sad principles for accurate perception of gestures.

Similarly, the table of confusion (Table VI) clearly show that the participants were not confident in predicting when the intended emotion was neutral. There were 16 instances of false negatives and 23 instances of true positives, which resulted in a sensitivity of only 58.97%. Because the TPR is less than 75% for the neutral intended emotion, we are not able to reject the null hypothesis. However, the participants were very confident predicting when the intended emotion was not neutral. There were 8 instances of false positives and 148 instances of true negatives, which resulted in a specificity of 94.87%.

TABLE V
CONFUSION MATRIX

	Happy (Predicted)	Neutral (Predicted)	Sad (Predicted)	Accuracy
Happy (Actual)	99	5	0	95.19%
Neutral (Actual)	10	23	6	58.97%
Sad (Actual)	0	3	49	94.23%

TABLE VI
ANALYSIS OF SENSITIVITY & SPECIFICITY

		Positive (Predicted)	Negative (Predicted)	Sensitivity/ Specificity
Happy	**Positive (Actual)**	99	5	95.19% (TPR)
	Negative (Actual)	10	81	89.01% (TNR)
Neutral	**Positive (Actual)**	23	16	58.97% (TPR)
	Negative (Actual)	8	148	94.87% (TNR)
Sad	**Positive (Actual)**	49	3	94.23% (TPR)
	Negative (Actual)	6	137	95.80% (TNR)

For both the happy and sad intended emotions, the participants were confident predicting both when the emotion was present and was not present. For happy, there were 5 instances of false negatives and 99 instances of true positives, which resulted in a sensitivity of 95.19%. This high TPR for the happy intended emotion allows us to reject the null hypothesis. There were 10 instances of false positives and 81 instances of true negatives, which resulted in a specificity of 89.01%. For sad, there were 3 instances of false negatives and 49 instances of true positives, which resulted in a sensitivity of 94.23%. This high TPR for the sad intended emotion allows us to reject the null hypothesis. There were 6 instances of false positives and 137 instances of true negatives, which resulted in a specificity of 95.80%.

Figure 3 illustrates that the participants were able to distinguish 7 gestures as extreme instances. G4, G6, G7, G13, and G14 were on average "very happy," whereas G11 and G12 were on average "very sad." In addition, the gesture with the smallest standard deviation of 0.376 was G9 with an average of 4.154 (sad). Once the range of happiness is combined into 1 category and the range of sadness is combined into 1 category, the participants completely agree on their perception of the gestures. In particular, G2, G4, G7, G9, G11, G13, and G14 have no deviation across all participants (standard deviation is 0). This suggests that implementing the aforementioned gestures into an actual social scenario would be ideal to enhance engagement and motivation.

Lastly, the movement of the upper body versus the lower body as discussed in [16] did not reveal any trends necessary for distinguishing extreme emotion (very happy/sad). All 15 of the gestures had some type of upper body movement, but 4 of the gestures incorporated lower body movement as well. Of the 4 gestures that incorporated lower body movement, 2 of them were classified as an extreme emotion (50%). However, of the 11 gestures that did not incorporate lower body movement, 5 of them were still classified as an extreme emotion (45.45%), while 6 of them were not classified as extreme emotion (54.54%).

VI. CONCLUSION

This study revealed that by altering head direction, arm direction, gesture size, and gesture speed on a humanoid robotic social agent, participants are able to achieve accurate perception when the intended emotion is happy or sad. By using these key principles to categorize the gestures, the standard deviation was kept consistently at a minimum when identifying emotion. In fact, 7 of the gestures yielded no standard deviation across all the participants, which is remarkable. When using this framework, the participants are very confident in identifying when the intended emotion is happy, not happy, sad, not sad, and not neutral. However, participants are not confident identifying when the intended emotion is neutral. Ultimately, this work suggests that engagement and motivation during social interaction can be optimized through the use of happy and sad gestures derived using the described framework.

VII. FUTURE WORK

In the near future, we plan to expand this study by applying the proposed framework to a variety of robotic platforms. This will allow us to tease out data associated with the robot and/or the gesture design and, ultimately, further generalize the model to increase its value. In addition, we would like to combine the evaluated happy and sad gestures with the evaluated verbal expressions in [4] to observe its impact in social interaction. Lastly, we will incorporate these gestures and verbal expressions into a learning scenario to evaluate how engagement, motivation, and academic performance are affected.

REFERENCES

[1] L. Brown, R. Kerwin, and A. Howard, "Applying behavioral strategies for student engagement using a robotic educational agent," in *IEEE SMC*, 2013.

[2] L. Brown and A. Howard, "Engaging children in math education using a socially interactive humanoid robot," in *IEEE-RAS Humanoids*, 2013.

[3] M. Saerbeck, T. Schut, C. Bartneck, and M. Janse, "Expressive robots in education: Varying the degree of social supportive behavior of a robotic tutor," in *CHI*, 2010, pp. 1613–1622.

[4] L. N. Brown and A. M. Howard, "The positive effects of verbal encouragement in mathematics education using a social robot," in *ISEC*, 2014.

[5] S. Garcia-Vergara, L. Brown, H. W. Park, and A. Howard, *Engaging Children in Play Therapy: The Coupling of Virtual Reality Games with Social Robotics*, 2014.

[6] D. Brooks and A. Howard, "Quantifying upper-arm rehabilitation metrics for children through interaction with a humanoid robot," in *Appl. Bionics Biomech.*, 2012.

[7] A. Howard, "Robots learn to play: Robots emerging role in pediatric therapy." in *International Florida Artificial Intelligence Research Society Conference*, 2013.

[8] H. W. Park and A. Howard, "Case-based reasoning for planning turn-taking strategy with a therapeutic robot playmate," in *IEEE International Conference on Biomedical Robotics and Biomechatronics*, 2010.

[9] R. Looije, M. A. Neerincx, and V. de Lange, "Children's responses and opinion on three bots that motivate, educate and play," *Journal of Physical Agents*, vol. 2, no. 2, June 2008.

[10] C. Kidd and C. Breazeal, "Robots at home: Understanding long-term human-robot interaction," in *IROS*, 2008, pp. 3230–3235.

[11] W. Taggart, S. Turkle, and C. D. Kidd, "An interactive robot in a nursing home," in *Toward Social Mechanisms of Android Science*, 2005.

[12] R. Tiberius and J. Billson, "The social context of teaching and learning," in *New Directions for Teaching and Learning*, 1991.

[13] J. A. Russell, "A circumplex model of affect," *Journal of Personality and Social Psychology*, vol. 39, pp. 1161–1178, 1980.

[14] A. Beck, L. Canamero, A. Hiolle, L. Damiano, P. Cosi, F. Tesser, and G. Sommavilla, "Interpretation of emotional body language displayed by a humanoid robot: A case study with children," *International Journal of Social Robotics*, 2013.

[15] J. Li and M. Chignell, "Communication of emotion in social robots through simple head and arm movements," *International Journal of Social Robotics*, 2013.

[16] E. Schegloff, "Body torque," *Soc Res*, 1998.

[17] M. J. Gielniak and A. L. Thomaz, "Enhancing interaction through exaggerated motion synthesis," in *HRI*, 2012.

[18] B. Burger and R. Bresin, "Communication of musical expression by means of mobile robot gestures," *Journal of Multimodal User Interfaces*, 2013.

[19] S. Dahl and A. Friberg, "Visual perception of expressiveness in musicians' body movement," 2007.

[20] K. Isbister, K. Hook, M. Sharp, and J. Laaksolahti, in *CHI*, 2006.

[21] I. Ha, Y. Tamura, H. Asama, J. Han, and D. Hong, "Development of open humanoid platform darwin-op," in *SICE*, 2011, pp. 2178–2181.